

Beitrag für	Handbuch Bibliothek. Metzler Verlag
Autor	Stefan Gradmann
Gliederungsziffer, Überschrift	5.5 Container – Content – Context: zur Evolution bibliothekarischer Metadaten von Katalogdaten zu Library Linked Data
Zugehörige Abb., Grafiken. Bildtexte	
	1
Status	Erstentwurf Gradmann vom 28.05.2012noch Zitat am Anfang prüfen: hubris / hybris
Umfang	SOLL: 34.400. IST: 29.800
An Verlag	

5.5 Container – Content – Context: zur Evolution bibliothekarischer Metadaten von Katalogdaten zu Library Linked Data

A world of exhaustive, reliable metadata would be a utopia. It's also a pipe-dream, founded on self-delusion, nerd hubris and hysterically inflated market opportunities. (Doctorow 2001)

5.5.1 Metadaten: Versuch einer Begriffsbestimmung

Der landläufige Versuch einer Bestimmung des Begriffes Metadaten beginnt meist mit deren Definition als „data about data“ (so etwa bei Doctorow 2001) – und setzt damit schon ein potentielles Missverständnis in die Welt, denn so, wie wir inzwischen „Daten“ zunehmend synonym mit „digitalen Daten“ begreifen, verstehen wir auch diese Definition verkürzend dahingehend, dass Metadaten digitale Daten über digitale Daten sind. Damit wird die 'Digitalität' zum impliziten Charakteristikum des Begriffs – und ist doch eigentlich eine eher ephemere Eigenschaft vieler (beileibe nicht aller!) Metadaten.

Tatsächlich sind Metadaten allgemein verstanden Aussagen über jede beliebige identifizierbare, benennbare, adressierbare oder in irgendeiner Weise handhabbare Entität unserer Welt – also das, was Berners-Lee et. al. (1998) als „resource“ bezeichnet. Diese weite Definition von Metadaten macht sie als Teil von Aussagenmodellen begreifbar und umfasst zugleich einige der in Doctorow (2001) polemisch angerissenen Dilemmata. In diesem Verständnis von Metadaten als Teile von Aussagenmodellen sind drei Konstituenten wesentlich: der Gegenstand der Aussage (ihr Thema), die Annahmen darüber, was über dies Thema ausgesagt werden kann (das Entitäten- oder Datenmodell) und die Kodierung und Serialisierung, die für die Bildung der Aussagen Verwendung findet.

5.5.1.1 Thema: der Aussagegegenstand

Der Gegenstand einer Aussage, das „Gegebene“ (so eine etymologienähe Übersetzung von „Data“) ist nur scheinbar eine Selbstverständlichkeit. Spätestens seit Kant wissen wir, dass uns der Blick auf die „wirkliche“ Natur der Dinge versperrt ist und wir nur Anschauungen kennen, uns mithin also niemals sicher sein können, dass ein anderer Mensch eine Entität in gleicher Weise wahrnimmt, wie wir es tun. Insofern ist die Definition von Berners-Lee et al. (1998), derzufolge „A resource can be anything that has identity“ eigentlich nur eine Problemverschiebung: Sie löst eben nicht die Frage, was Identität eigentlich konstituiert, sondern verschreibt sich im Grunde einem relativ platten Nominalismus (wenngleich auf hohem technischen Niveau). Dennoch will ich für die Zwecke dieser Darstellung einmal annehmen, dass wir uns in vielen Fällen darauf einigen können, was die Identität dessen ist, worüber wir Aussagen machen wollen.

5.5.1.2 Grammatik: das Datenmodell

Wenn wir uns nun über die Bezugsentität verständigt haben, so ist als nächstes zu klären, nach Maßgabe welches Modells man Aussagen über sie machen kann – und dies ist sehr weitgehend von der Modellierungsintention abhängig. So ist etwa die unter <http://www.kfs.org/~jonathan/witt/tlph.html> verfügbare Ausgabe von Wittgensteins 'Tractatus' für einen Textwissenschaftler bzw. einen Corpuslinguisten bzw. einen Philosophen etwas je völlig anderes, weil jeweils ein ganz unterschiedliches Wahrnehmungs- und Beschreibungsmodell unterlegt wird. Im Datenmodell des Textwissenschaftlers gibt es 'Absätze', 'Sätze', 'Fußnoten' und andere Teilentitäten, die sich mit dem Dokumentmodell der Text Encoding Initiative (TEI) beschreiben lassen, für den Linguisten geht es um 'Sätze' und 'Wörter' und deren Analyse in Begriffen eines Sprachsystems und der sprachlichen Pragmatik – und für den Philosophen liegt ein System abstrakter Aussagen mit logischen Beziehungen der Aussagen untereinander vor: Dementsprechend wird jeder Wissenschaftler dies Objekt mit einer anderen grammatischen Abstraktionsform beschreiben, die mit den Kategorien korreliert, in denen er seine 'Welt' konstituiert.

5.5.1.3 Syntax: Kodierung und Serialisierung

Schließlich kann ein und dieselbe Aussage völlig unterschiedlich kodiert und serialisiert werden: Sie kann in das zu beschreibende Objekt als 'Tagging' eingebettet oder als eigenständiger Datensatz mit einem Zeiger auf die beschriebene Entität auftreten. Sie kann weitgehend natürlichsprachig oder stark formalisiert (etwa in XML) verfasst sein. Und schließlich kann die verwendete Syntax für die platzsparende Nutzung eines Katalogzettels mit normierter Größe optimiert und dem entsprechend in weiten Teilen durch Abkürzungsregeln konstituiert sein.

5.5.1.4 Ein Beispiel: RDF

Im Lichte dieses Versuchs einer Begriffsbestimmung ist wahrscheinlich das Resource Description Framework (RDF) des W3C der reinste Prototyp eines Metadatenmodells: Im Kern sieht es einfache, dreiwertige Satztypen vor, die „Triples“, die aus einem Subjekt (= Aussagegegenstand), einem Prädikat (= einem Attribut des Gegenstandes) und einem Objekt (= dem Wert, den dieses Attribut annimmt) bestehen. Mit solchen Tripeln lassen sich typisierte Verbindungen zwischen zwei 'resources' herstellen – wie zum Beispiel: „Arno Schmidt“ – istAutorVon – „Zettels Traum“ – oder Aussagen über den Zustand eines Gegenstandsattributs machen (wie zum Beispiel „Zettels Traum“ – hatGewicht – 10).

In der für den gewählten Modellierungsraum maßgeblichen, als RDF-Schema (oder in der Ontology Working Language [OWL] als Spezialisierung von RDF-Schema) verfassten Grammatik wäre dann etwa für das erste Beispiel formuliert, dass eine Instanz der Klasse „autor“ eine natürliche Person männlichen Geschlechts sein muss bzw. für das 2. Beispiel, dass der Literal im Eigenschaftswert immer eine natürliche Zahl in der Maßeinheit Kilogramm sein soll.

5.5.2 Bibliothekarische Metadaten

Metadata (a fancy name for an inferior form of cataloguing)
(Gorman 2000)

Im Vergleich mit einem so hochgradig generischen Aussagenmodell wie RDF waren bibliothekarische Metadaten bis vor ganz kurzer Zeit thematisch weitestgehend fokussiert auf in sich geschlossene, monolithische Informationsobjekte vom Typ 'Buch', wobei über die letzten 2000 Jahre eine Entwicklung vor allem hinsichtlich des Objektmodells stattgefunden hat. Thematisch war jedoch das 'Buch' vor allem nach der Erfindung des Buchdrucks und der damit entstehenden Gutenberg-Galaxis ein ungleich stabileres Objekt als etwa die ungleich diverseren Beschreibungsgegenstände etwa in Museen, in denen von einer Tonscherbe bis zu Multimediaobjekten eine ungleich größere Vielzahl von Objekttypen modelliert werden muss – was die relative Komplexität und Abstraktheit eines museumsbezogenen Metadatenstandards wie das Conceptual Reference Model (CIDOC CRM) erklärt.

5.5.2.1 Traditionelle Katalogdaten

Schon die ersten uns bekannten Bibliothekskataloge – wie etwa die Πίνακες der Bibliothek von Alexandria – verzeichneten demgegenüber relativ uniforme Informationsobjekte: Schriftrollen (oder später eben Bücher), zweidimensional und sequentiell organisiert und beschriftet, gelegentlich illustriert – aber doch ganz weitgehend am Textparadigma orientiert. Immerhin war aber Kallimachos als Direktor der Bibliothek von Alexandria noch zugleich Wissenschaftler und Dichter: die so entstehenden Metadaten waren also noch stark inhaltsbezogene philologische Traktate. Und noch im Mittelalter waren die Literatur- und die Metadatenproduktion bei weitem nicht so entkoppelt wie in späteren Zeiten: Im Bauplan des imaginären Klosters von St. Gallen liegen Bibliothek und Scriptorium noch direkt nebeneinander im dessen Zentrum.

Dies änderte sich grundlegend mit der Erfindung des Buchdrucks und mit der Öffnung der Gutenberg-Paranthesis: Autoren-/Verlagssicht und Bibliotheksperspektive wurden nun disjunkt und damit die Katalogisierungspraxis in den allermeisten wissenschaftlichen Bibliotheken zunehmend containerbezogen und nicht mehr von der Auseinandersetzung mit den Inhalten geprägt. Die nun entstehenden Metadaten erfassten mit großer Akribie Attribute des Titelblattes, Seitenzahl und Einbandart – kaum jedoch inhaltliche Aspekte des Informationsobjektes. Entsprechende Schwerpunkte setzten auch die nun entstehenden bibliothekarischen Regelwerke für die Metadatengenerierung: Sie berücksichtigten (wie etwa die „Preußischen Instruktionen“) den Inhalt des beschriebenen Buches überhaupt nicht oder relegierten ihn – wie im Falle der Regeln für die „Alphabetische Katalogisierung“ (RAK) in das Schwesterregelwerk „Regeln für die Schlagwortkatalogisierung“ (RSWK).

Hinzu kam, dass die bibliothekarischen Regelwerke neben dem containerbezogenen Beschreibungsmodell zwei weiteren Restriktionen genügen mussten: Die Beschreibungsinformation sollte auf Katalogzetteln in standardisierter Größe passen und diese Katalogzettel mussten in eine sequentielle Ordnung gebracht werden. Letzteres führte

dazu, dass ein Regelwerk wie die „Preußischen Instruktionen“ einen erheblichen Aufwand betrieben, um eine als grammatikalisch logisch postulierte sequentielle Anordnung der Buchtitel realisierbar zu machen, und ersteres führte mit zur Definition der „International Standard Bibliographic Description“ (ISBD), mit der die Beschreibungsinformation zu einer kompakten Zeichenkette mit komplex formalisierter Syntax wurde.

5.5.2.2 Katalogautomation

Daran änderte sich auch in der in den 60er Jahren des letzten Jahrhunderts beginnenden Phase der IT-gestützten Katalogautomation im ersten Angang wenig: Die Fixiertheit auf Containerobjekte mitsamt dem Zwang, sequentiell organisierte Katalogmetadaten zu produzieren, und der im neuen Umfeld eigentlich funktionslosen ISBD führte dazu, dass in der damals definierten Syntax des bibliografischen Datenformats MARC sogar die Interpunktion der ISBD zum Feldinhalt wurde, wie die eingekreisten Elemente in Abbildung 1 zeigen:

```
00882cam 2200253 a 450400100080000000800410000801000130004
90200015000620350013000770400018000900490009001080500022001
17082001900139100001800158245012100176260004900297300002500
34644000500037150400510042165000640047265000470053670000190
0583700002600602|3230053|920608s1992 nyu bf 001 0
eng | $a92021087| $a1555701205| $a26132811| $aDLC$cDL
C$dAGL| $aHNKA|00$aZ678.9$b.C58 1992|00$a025/.00285$220|1
$aCohn, John M.|10$aPlanning for automation :$ba how-to-do
it manual for librarians /$cJohn M. Cohn, Ann L. Kelsey, Ke
ith Michael Fiels.| $aNew York :$bNeal Schuman Publishers,
$c1992.| $avi, 116 p. :$c28 cm.| 0$aHow-to-do it manuals
for libraries ;$vno. 25.| $aIncludes bibliographical refer
ences and index.| 0$aLibraries$xAutomation$xManagement$xHan
dbooks, manuals, etc.| 0$aLibrary planning$xHandbooks, manu
als, etc.|10$aKelsey, Ann L.|10$aFiels, Keith Michael.*
```

Die Katalogautomation als Teil der nun entstehenden Systeme für die Bibliotheksau-
tomation änderte mithin am überkommenen containerbezogenen Katalog-Paradigma
grundlegend nicht viel und emulierte dessen Funktionsweise letztlich nur im digitalen
Umfeld.

5.5.2.3 FRBR

Eine erste wirkliche Veränderung brachte in dieser Hinsicht die IFLA-Studie mit dem
Titel „Functional Requirements for Bibliographic Records“ (IFLA 1998) Vor allem die
dort als Teil der Gruppe I der beschreibungsrelevanten Einheiten vorgenommene Un-
terscheidung der vier Ebenen *Werk – Expression – Manifestation* und *Exemplar* (work,
expression, manifestation und item im Original) beginnt eine Gewichtsverschiebung
hin zum Inhalt der Informationsobjekte. Dabei sind die Einheiten Werk und Expression
entscheidend und eindeutig inhaltlich konstituiert, die materiellen Einheiten Manifesta-
tion und Exemplar werden zu nicht primär beschreibungskonstitutiven Derivatformen.
Es ist anzunehmen, dass vor allem diese fundamentale Perspektivverschiebung die
Ursache dafür war, dass FRBR einige Jahre lang innerhalb der Bibliotheks-Community
eher ein Fremdkörper blieb und eine praktische Konkretisierung bis auf vereinzelte
Experimente wie bei OCLC im FictionFinder -Projekt (OCLC 2009) unterblieb.

5.5.2.4 RDA

Dies ändert sich grundlegend mit dem neuen Regelwerk Resource Description and Access (Joint Steering Committee for Development of RDA 2010), dessen Kern eben die Terminologie und Perspektive der Gruppe I der FRBR darstellt. Damit wird erstmals eine inhaltlich fundierte Entscheidung darüber erforderlich, in welchem Zusammenhang zu beschreibende Informationsobjekte zueinander stehen: Eine deutsche Übersetzung von Shakespeares „Hamlet“ wird nun zur *Expression* dieses *Werks*, und das Buch mit der ISBN 978-3-86057-567-3 zu einer *Manifestation*, von der Bibliotheken einzelne *Exemplare* besitzen. Für die Katalogisierungsarbeit konstitutiv sind dabei aber nun weniger die Container-Eigenschaften der beiden letzten Ebenen als vielmehr inhaltliche Zusammenhänge und Abhängigkeiten – und dies verleiht RDA bei allen Unfertigkeiten und rückwärtsgewandten Kompromissen letztlich doch eine gewisse Sprengkraft, die vor allem die nicht enden wollenden Diskussionen über die Einführung von RDA zu befeuern scheint – mehr dazu am Schluss dieses Beitrages.

5.5.2.5 Syntaxen

Schon die „Preußischen Instruktionen“ enthielten Serialisierungsregeln, welche die Reihenfolge der Elemente der Titelaufnahme und die zu ihrer Trennung verwendete Interpunktion betrafen. Die ISBD enthielt dann den letzten Serialisierungsansatz dieser Art als Formalisierung einer komplexen Zeichenkette.

Mit dem Aufkommen der IT-gestützten Katalogautomatisierung änderte sich dies: Die Beschreibungsentität war nun ein Datensatz, ein 'record', und dessen interne Syntax war fortan als eine Folge von Attributen mit meist numerischen Feldbezeichnern und Feldinhalten als Freitext oder weiter durch Unterfeldbezeichner strukturiert serialisiert. Dies Grundprinzip war dann in der Folge allen Formaten für die Serialisierung von Kataloginformationen gemeinsam, die in erster Linie entstanden, um diese Informationen über Anwendungs- und Datenbankgrenzen hinaus austauschbar zu machen. Die wichtigsten dieser Austauschformate (oder präziser gesprochen Formatfamilien) waren das MARC-Format (in seiner letzten Fassung als MARC 21 „Format for Bibliographic Data“) und UNIMARC. Deutschland war auch hier mit der Entwicklung des Formates MAB (Maschinenlesbares Austauschformat für Bibliotheken) einen Sonderweg gegangen, den es aber nunmehr am Ende dieser Epoche mit der Übernahme von MARC 21 verlassen hat. Auch wenn der Aufbau der 'Records' teilweise technisch unterschiedlich realisiert war (und sich wie im Fall des MARC-Headers auch den Erfordernissen und Restriktionen der maschinellen Verarbeitung verdankte), war doch das Grundprinzip dieser Serialisierungsformen immer noch dasjenige der vom beschriebenen Objekt abgesetzten Katalogdatensätze mit einem 'Zeiger' auf das beschriebene Objekt – sei es eine Katalogsignatur oder eine Identnummernverknüpfung in einem Datenbanksystem.

5.5.3 Nicht-bibliothekarische bibliographische Metadaten

Inzwischen war jedoch außerhalb der Bibliothekswelt eine wichtige Veränderung eingetreten: Beginnend mit Tim Berners-Lees berühmtem 'proposal' für das Informationsmanagement zuerst am CERN und in der Hochenergiephysik war seit Beginn der 1990er Jahre das World Wide Web entstanden, in dem zunehmend auch Informationsressourcen zu speichern, zu referenzieren und zu beschreiben waren. Dabei entstanden nunmehr unabhängig von der bibliothekarischen Katalogisierung Metadatenstandards für die Beschreibung und das Retrieval von WWW-Informationsressourcen, als deren prominentester Vertreter das seit 1995 existierende Set der 15 Attribute von Dublin Core (DC) gelten darf. DC wurde dann in der Folge im Abstract Model der Dub-

lin Core Metadata Initiative (DCMI) soweit generalisiert und verfeinert, dass es als Metadatenformat für die Wissensrepräsentation gesehen werden kann (s. u.)

Weiter entstanden nach dem Aufkommen und der erfolgreichen Standardisierung von XML auf dieser Syntax basierte Metadatenstandards, die nunmehr aber zunehmend in die beschriebenen Objekte eingebettet waren und der Modellierung der Objekt-Binnenstrukturen dienten. Beispiele sind die Encoded Archival Description (EAD, für Findbuchstrukturen in Archiven) der die Guidelines der Text Encoding Initiative (TEI) für beliebige print-analoge Dokumentstrukturen.

5.5.4 Repräsentation von Inhalten und Wissensrepräsentation

Mit dem beginnenden Aufbau des 'Semantic Web' gewann dann seit den frühen 2000er Jahren das oben schon dargestellte Aussagensystem RDF stetig an Bedeutung. Dies zum einen als Grundlage für höhere, auf die Unterstützung von Schlussystemen hin optimierte Sprachkonstrukte wie die Ontology Working Language (OWL) – aber auch als Mittel, Aussagen über beliebige im WWW repräsentierte Entitäten der 'wirklichen' Welt Aussagen machen zu können, womit das Web gleich eine zweifache Erweiterung erfuhr: Eine syntaktische Erweiterung um die Möglichkeit, Verbindungen zwischen Web-Ressourcen zu typisieren und eine Erweiterung im Erstreckungsbereich, nach der so ziemlich alles in der Welt im „Web der Dinge“ repräsentierbar wurde.

Damit wurde aus dem Web der Dokumente (denn als solches war das WWW von Berners-Lee ursprünglich angelegt) ein gigantischer Raum für die Generierung und Organisation von Wissen im Sinne kontextualisierter Information. Dabei hielt die 'Community' des WWW stets Türen für die Bibliothekswelt offen. Hatte schon Bray bezogen auf den ersten RDF-Standard von 1998 als erstes Einsatzszenario für die von ihm so bezeichneten „Divine Metadata for the Web“ die Bibliothek benannt, so ist die Positionierung des Simple Knowledge Organisation System (SKOS) durch das World Wide Web Consortium in dieser Hinsicht kaum deutlicher denkbar. Als Ziel des Standards wird dort formuliert: „making the accumulated experience and wisdom of knowledge organization in the library and information sciences accessible, applicable and transferable to the technological context of the Semantic Web“ (World Wide Web Consortium 2009).

Das hiermit formulierte Angebot beginnen die Bibliotheken inzwischen auch systematisch anzunehmen: Die Normdaten der Deutschen Nationalbibliothek oder auch die 'Subject Headings' der Library of Congress sind mittlerweile als SKOS-Daten im WWW verfügbar.

5.5.5 Die Gutenberg-Klammer schließt sich: beginnende Auflösung der monolithischen Containerformate

Parallel zur Entwicklung im Semantic Web deutet sich eine zweite grundlegende Entwicklung an, welche die von Bibliotheken zu erschließenden Informationsobjekte selbst betrifft. Die Gutenberg Paranthesis Group formuliert dies folgendermaßen: „From the vantage point of the early 21st century it seems likely that conventional notions of the text, which since Gutenberg have often been conflated with the book, will be radically transformed.“ Und schon Pédauque (2003) war ein systematischer Versuch, den Dekonstruktionsprozess zu modellieren, dem die jahrhundertealten vergleichsweise monolithisch und linear angelegten Dokumentformate momentan unterliegen, die als in sich geschlossene Container über lange Zeit die bibliothekarische Katalogisierungspraxis in ihrem Wesenskern geprägt hatten.

In dem Maße nun wie diese Container sich in kleinere Einheiten aufzulösen beginnen und – wie etwa in dem von Barend Mons und Jan Velterop propagierten Ansatz des 'Nano-Publishing' – im Einzelfall schon durch Aggregationen von RDF-Tripeln ersetzt werden, verliert natürlich auch das traditionelle Katalogparadigma der Bibliotheken seine Bezugspunkte.

Dies schlägt sich in neuen und neuartigen Metadatenmodellen nieder: Auf die Atomisierung und Diversifizierung der Containerstrukturen reagieren Metadatenstandards wie das Metadata Object Description Schema (MODS) oder auch der Metadata Encoding and Transmission Standard (METS), die zunehmend Informationsobjekte nicht mehr als 'Black Box' modellieren, sondern die Mikrostrukturen komplexer, zusammengesetzter digitaler Informationsobjekte mitsamt deren Semantik in den Blick rücken.

5.5.6 Linked Data: 'Context is King'

Mit dieser Tendenz zur Auflösung monolithischer Dokumentstrukturen einher geht jedoch eine weitere Entwicklung, die sich als Symptom letztlich auch in der momentan omnipräsenten Diskussion über Forschungsprimärdaten und deren Verortung im Bibliotheksumfeld niederschlägt: Hier wird dann nämlich tendenziell auch die klare Unterscheidung von Objektdaten und Metadaten obsolet. Konstitutiv für eine 'Dokumenteinheit' sind nun nicht mehr die klaren und scheinbar intuitiv fassbaren Grenzen eines Bucheinbandes, sondern letztlich nurmehr dessen Kontextbezüge, die zudem solche scharfen und scheinbar intuitiv fassbaren Grenzziehungen nicht mehr wirklich erlauben.

Wo in Abbildung 2 die Grenze zwischen den Tripel-Aggregationen A und B verläuft und wie weit die zwischen ihnen liegenden Kontextbezüge dieser beiden 'Dokumente' einander überlappen ist eine nur noch schwer zu beantwortende Frage.

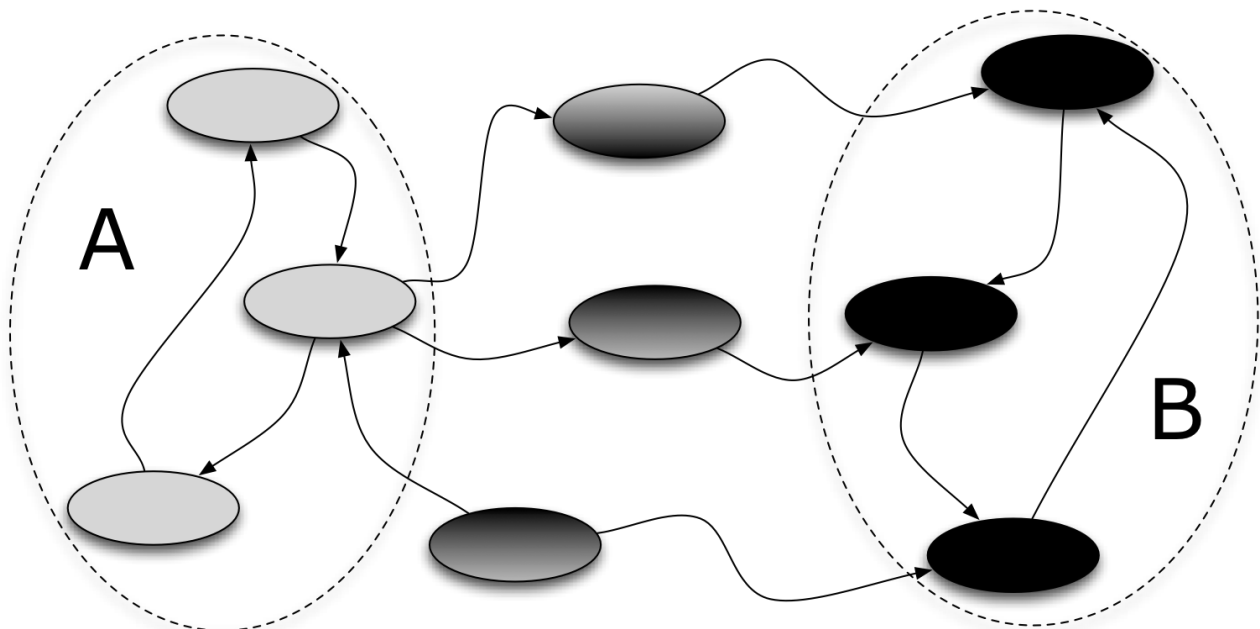


Abbildung 2: Zwei 'Dokumente' mit überlappendem Kontext

So werden denn letztlich in diesem Netz der Linked Data Objektattribute und die Objektidentität selbst zunehmend nur noch aus Strukturbeziehungen im Aggregationskontext erschließbar und sind zudem als nie vollständig stabil zu denken. In diesem

Umfeld liegt der Fokus zunehmend auf Kontextualisierung und der Arbeit mit Kontextbezügen, Figuren der Identität werden zunehmend ersetzt von solchen der 'Différance' im Sinne Derridas.

5.5.7 Die Zukunft der bibliothekarischen Metadaten

Mit der Ankündigung einer Working Group of the Future of Bibliographic Control hat die Library of Congress im Jahr 2011 faktisch das Ende der exklusiven bibliothekarischen Katalogisierungs-nische eingeläutet. Das Ziel ist nun nicht mehr die Fortschreibung des bibliothekarischen Regelbiotops unter veränderten technischen Randbedingungen. Programmatisch heisst es dort vielmehr: „The new bibliographic framework project will be focused on the Web environment, Linked Data principles and mechanisms, and the Resource Description Framework (RDF) as a basic data model.“ (Library of Congress 2011). Damit verabschiedet sich die weltgrößte Bibliothek nicht nur vom Serialisierungsformat MARC, sondern postuliert auch zumindest implizit einen Kulturbruch: Hatten Bibliotheken bis dahin die Gestaltung von Katalogisierungsregelwerken wie auch die Katalogisierungspraxis selbst immer als ihre ureigene Domäne gesehen, wird nun der Exklusivitätsanspruch in beiden Bereichen aufgegeben zugunsten einer Einordnung in den größeren Kontext der Technologien und Modellierungsansätze des WWW und insbesondere des Linked Data Web.

Damit beginnt die Bibliothekswelt eine weitreichende Neuorientierung in der von der Library Linked Data Incubator Group (World Wide Web Consortium 2011) empfohlenen Richtung. Der dabei inzwischen erreichte beeindruckende Stand ist in einem weiteren Report der 'Incubator Group' (World Wide Web Consortium 2011b) zusammengefasst.

Damit scheint sich die Waagschale zu neigen in einer Entscheidungssituation grundlegenden Zuschnitts. Diese hatte der Wissenschaftsrat im Jahr 2011 bezogen auf die Situation des deutschen Verbundsystems folgendermaßen gekennzeichnet: „entweder eine konsequente Integration dieser Daten auf der technischen Basis von OCLCs WorldCat oder aber in Form von in das offene Web integrierten Linked Open Data“ (Wissenschaftsrat 2011, 40).

Mit der sich nun weltweit abzeichnenden Entwicklung hin zu auf Linked Data basierenden Ansätzen haben die Bibliotheken einen weiten Weg zurückgelegt: Aus der Zeit der fast exklusiv containerbezogenen Bestandskataloge über eine dann einsetzende (Wieder-)Entdeckung der Dokumentinhalte hin zu einer zunehmenden Fundierung des eigenen Tätigkeitsprofils in den Kontextualisierungstechniken des Linked Data Web.

Damit sind nicht etwa alle denkbaren Probleme gelöst. Zwar eröffnet sich mit dieser Entwicklung das Potential einer weitgehend redundanzfreien Modellierung bibliographischer Daten weltweit: Jede Entität müsste in zukünftigen Architekturmodellen für 'Linked Library Metadata' auf den Ebenen Werk, Expression und Manifestation weltweit eigentlich nur noch einmal als RDF-Aggregation beschrieben sein, damit Bibliotheken dann in Verlinkung auf diese gemeinsame Verzeichnungsebene nur noch mit ihren bestandsspezifischen Exemplardaten befasst sind. Damit werden zumindest theoretisch beträchtliche Arbeitszeitpotentiale für andere Aufgaben verfügbar und zugleich Architekturmodelle für die Bibliotheksautomation denkbar, welche die wenigen verbleibenden Hersteller solcher Systeme zum grundlegenden Nachdenken bringen müssen – faktisch geschieht dies auch schon zumindest bei den Herstellern ExLibris und OCLC, die zunehmend über Cloud-basierte Servicemodelle nachdenken. Selbst die OCLC-Kooperative, deren Geschäftsmodell im Kern bis dahin der Zugang zur weltweit größten geschlossenen bibliothekarischen Metadatenressource in Gestalt des WorldCat war, setzt inzwischen nicht mehr wirklich auf die Nachhaltigkeit dieser strategischen Aus-

richtung.

Zugleich wird dann aber die Bibliothekswelt die neuen und nicht unerheblichen Probleme der Linked Data Community teilen: Fragen der Versionierbarkeit, der Autorisierungsszenarien, der Notation von Provenienzinformatoren etwa sind im Linked Data Umfeld noch weitgehend ungelöst, und auch die Qualität vieler Linked Data Ressourcen ist mindestens verbesserungsfähig, so wie auch der Umgang mit den vielen (meist partiellen) semantischen Redundanzen in den Metadaten der Linked Data Community noch eine Menge Arbeit nach sich ziehen wird.

Hier können gerade Bibliothekare mit ihren vergleichsweise hochwertigen Metadaten und ihrer professionellen Disziplin beim Umgang mit Metadaten wertvolles Potenzial einbringen. Insofern muss der bibliothekarischen Zunft nicht wirklich bang sein: Gerade weil das 'semantische' Web der Linked Data ohne Metadaten im Grundansatz nicht funktionieren kann, sind die Chancen gut für Bibliothekare – solange sie eine regressiv-defensive Grundhaltung vermeiden, wie sie aus dem oben als Kapitelmotto verwendeten Zitat von Gorman spricht. In die abgeschotteten Regelwerksbiotope der traditionellen Bestandskataloge führt wohl kein Weg zurück! Metadaten hingegen werden auf absehbare Zeit ein unverzichtbarer Bestandteil auch avancierter Informations- und Wissensarchitekturen sein.

5.5.8 Literatur

- AACR2 (2005). Joint Steering Committee for Revision of AACR: Anglo-American Cataloguing Rules Second Edition, 2002 Revision: 2005 Update
- Berners-Lee, T., Fielding, R., & Masinter, L. (1998). Uniform Resource Identifiers (URI): Generic Syntax. IETF = <http://www.ietf.org/rfc/rfc2396.txt>
- Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The Semantic Web. *Scientific American*, 284(5), 34-43.
<http://www.nature.com/doifinder/10.1038/scientificamerican0501-34>
- Bray, T. (2001). What is RDF? = <http://www.xml.com/pub/a/2001/01/24/rdf1.html?page=1>
- CIDOC CRM (2011). The CIDOC Conceptual Reference Model. = http://www.cidoc-crm.org/official_release_cidoc.html
- Doctorow, C. (2001). Metacrap: Putting the torch to seven straw-men of the meta-utopia = <http://www.well.com/~doctorow/metacrap.htm>
- Gorman, M. (2000). From Card Catalogues to Webpacs: Celebrating Cataloguing in the 20th Century. Presented at: Library of Congress Bicentennial Conference on Bibliographic Control for the New Millennium Washington, D.C., November 15th 2000 = http://loc.gov/catdir/bibcontrol/gorman_paper.html
- Gradmann, S. (1998). Cataloguing vs. Metadata. Old wine in new bottles? *Proceedings of 64th IFLA General Conference August 16 - August 21, 1998 in Amsterdam (Netherlands)* = <http://archive.ifla.org/IV/ifla64/007-126e.htm>
- Greenberg, J. (2012). Metadata and Digital Information Metadata and Digital Information. *Encyclopedia of Library and Information Sciences*. S. 3610-3623.= <http://dx.doi.org/10.1081/E-ELIS3-120044415>
- Gutenberg Paranthesis Research Group / University of Southern Denmark (2010): Position Paper = http://www.sdu.dk/en/Om_SDU/Institutter_centre/Ilkm/Forskning/Forskningsproj

ekter/Gutenberg_projekt/PositionPaper

- IFLA (1998). Functional requirements for bibliographic records : final report / IFLA Study Group on the Functional Requirements for Bibliographic Records. — München : K.G. Saur, 1998. (UBCIM publications ; new series, vol. 19).
- Joint Steering Committee for Development of RDA (2010). RDA: Resource Description and Access = <http://www.rda-jsc.org/rda.html>
- Library of Congress. Metadata Object Description Schema (MODS). = <http://www.loc.gov/standards/mods/>
- Library of Congress. Metadata Encoding and Transmission Standard. = <http://www.loc.gov/standards/mets/>
- Library of Congress. (2011). A Bibliographic Framework for the Digital Age. = <http://www.loc.gov/marc/transition/news/framework-103111.html>
- Mons, B., & Velterop, J. (2009). Nano-Publication in the e-science era. Bioinformatics, 14. = http://www.nbic.nl/uploads/media/Nano-Publication_BarendMons-JanVelterop.pdf
- OCLC (2009). FictionFinder: A FRBR-based Prototype for Fiction in WorldCat = <http://www.oclc.org/research/activities/fictionfinder/default.htm>
- Pédauque, R. T. (2003). Document: Form, Sign and Medium, As Reformulated for Electronic Documents.
- PI (1975). Instruktionen für die alphabetischen Kataloge der preussischen Bibliotheken vom 10. Mai 1899. - 2. Ausg. in der Fassung vom 10. August 1908. Unveränderter Nachdruck. Wiesbaden: Harrassowitz.
- RAK (1983). Regeln für wissenschaftliche Bibliotheken: RAK-WB. Wiesbaden: Reichert.
- RSWK (1986). Regeln für den Schlagwortkatalog: RSWK. Berlin: DBI.
- Smiraglia, R. P. (Ed.). (2005). Metadata: A Cataloger's Primer. Binghamton, NY: Haworth Information Press.
- Sowa, J. (2000). Ontology, Metadata, and Semiotics. In B. Ganter & G. Mineau (Eds.), *Conceptual Structures Logical Linguistic and Computational Issues* (Vol. 1867, S. 55-81). Springer Berlin / Heidelberg. doi:10.1007/10722280_5
- Text Encoding Initiative. (2007). TEI: P5 Guidelines. Retrieved June 4, 2012, from <http://www.tei-c.org/Guidelines/P5/>
- Wissenschaftsrat. (2011). Empfehlungen zur Zukunft Verbundsystems in Deutschland. Köln. = <http://www.wissenschaftsrat.de/download/archiv/10463-11.pdf>
- World Wide Web Consortium (2009). SKOS Simple Knowledge Organization System Reference. W3C Recommendation 18 August 2009. = <http://www.w3.org/TR/skos-reference/>
- World Wide Web Consortium (2011a). Library Linked Data Incubator Group Final Report. Retrieved June 4, 2012, from <http://www.w3.org/2005/Incubator/llid/XGR-llid-20111025/>
- World Wide Web Consortium (2011b). Library Linked Data Incubator Group: Datasets, Value Vocabularies, and Metadata Element Sets. = <http://www.w3.org/2005/Incubator/llid/XGR-llid-vocabdataset-20111025/>

\$KStefan Gradmann\$K\$